

# REGULARITY IN MONGE'S MASS TRANSFER PROBLEM

QI-RUI LI, FILIPPO SANTAMBROGIO, AND XU-JIA WANG

**ABSTRACT.** In this paper, we study the regularity of optimal mappings in Monge's mass transfer problem. Using the approximation to Monge's cost function  $c(x, y) = |x - y|$  through the costs  $c_\varepsilon(x, y) = \sqrt{\varepsilon^2 + |x - y|^2}$ , we consider the optimal mappings  $T_\varepsilon$  for these costs, and we prove that the eigenvalues of the Jacobian matrix  $DT_\varepsilon$ , which are all positive, are locally uniformly bounded. By an example we prove that  $T_\varepsilon$  is in general not uniformly Lipschitz continuous as  $\varepsilon \rightarrow 0$ , even if the mass distributions are positive and smooth, and the domains are  $c$ -convex.

## 1. INTRODUCTION

The Monge mass transfer problem consists in finding an optimal mapping from one mass distribution to another one such that the total cost is minimized among all measure preserving mappings. This problem was first proposed by Monge [27] and has been studied by many authors in the last two hundred years: among the main achievements in the 20th century we cite [21] and [16].

In Monge's problem, the cost of moving a mass from point  $x$  to point  $y$  is proportional to the distance  $|x - y|$ , namely the cost function is given by

$$(1.1) \quad c_0(x, y) = |x - y|.$$

This is a natural cost function. In the last two decades, due to a range of applications, the optimal transportation for more general cost functions has been a subject of extensive studies. In order to present the framework more precisely, let  $\Omega$  and  $\Omega^*$  be two bounded domains in the Euclidean space  $\mathbb{R}^n$ , and let  $f$  and  $g$  be two densities in  $\Omega$  and  $\Omega^*$  respectively, satisfying the mass balance condition

$$(1.2) \quad \int_{\Omega} f(x) dx = \int_{\Omega^*} g(y) dy.$$

Let  $c$  be a smooth cost function defined on  $\Omega \times \Omega^*$ .

---

2010 *Mathematics Subject Classification.* 35J60, 35B65, 49Q20.

*Key words and phrases.* Optimal transportation, Regularity.

The first author was also supported by the Chinese Scholarship Council, the second by the French ANR Project ISOTACE : ANR-12-MONU-0013, the third by the Australian Research Council.

The problem consists in finding a map  $T : \Omega \rightarrow \Omega^*$  which solves

$$\min \int c(x, T(x)) f(x) dx \quad : \quad T_{\#} f = g,$$

where the last condition reads “the image measure of  $f$  through  $T$  is  $g$ ” and means  $\int_A g(y) dy = \int_{T^{-1}(A)} f(x) dx$  for all subsets  $A \subset \Omega^*$ .

The existence and uniqueness of optimal mappings were obtained in [4, 7, 20] if the cost function  $c$  satisfies

(A)  $\forall (x_0, y_0) \in \Omega \times \Omega^*$ , the mappings  $x \in \overline{\Omega} \rightarrow D_y c(x, y_0)$  and  $y \in \overline{\Omega}^* \rightarrow D_x c(x_0, y)$  are diffeomorphisms onto their ranges.

The regularity of optimal mappings was a more complicated issue. For the quadratic cost function, it reduces to the regularity of the standard Monge-Ampere equation, of which the regularity has been studied by many authors (see for instance [5, 6]). For general costs, the regularity was obtained in [26] if the domains satisfy a certain convexity condition,  $f, g$  are positive and smooth, and the cost function  $c$  satisfies the following structure condition

(B)  $\forall x \in \overline{\Omega}, y \in \overline{\Omega}^*$ , and vectors  $\xi, \eta \in \mathbb{R}^n$  with  $\xi \perp \eta$ ,

$$\sum_{i,j,k,l,p,q,r,s} \xi_i \xi_j \eta_k \eta_l (c_{ij,rs} - c^{p,q} c_{ij,p} c_{q,rs}) c^{r,k} c^{s,l}(x, y) \geq \beta_0 |\xi|^2 |\eta|^2,$$

where  $\beta_0$  is a positive constant. Loeper [24] showed that the optimal mapping may not be continuous if the condition (B) is violated, i.e. when there exist  $\xi, \eta \in \mathbb{R}^n$  with  $\xi \perp \eta$  such that the left hand side is negative. There are many follow-up researches on the regularity, in both the Euclidean space [23, 33] and on manifolds [2, 12, 18, 22, 25]. See also [31] for recent development.

Monge’s mass transfer problem is a prototype of the optimal transportation and the function (1.1) is the natural cost function. Therefore the existence and regularity of optimal mappings for Monge’s problem are of particular interest. However this cost function does not satisfy both key conditions, namely the condition (A) for the existence and the condition (B) for the a priori estimates.

The existence of optimal mappings for Monge’s problem has been studied by many researchers and was finally proved in [8, 32]. Prior to that, the existence was also obtained in [16] under some assumptions, and obtained in [30], with a gap fixed in [1]. See also [3, 9, 10] for the existence of optimal mappings when the norm (1.1) is replaced by a more general norm in the Euclidean space. The proofs in [8, 32] are very similar: both use the approximation  $|x - y|^{1+\varepsilon} \rightarrow |x - y|$  ( $\varepsilon \rightarrow 0$ ). The key point is choosing an approximation with strictly convex costs of the difference  $x - y$ , which satisfy the

assumption **(A)**. The optimal mapping for Monge's problem is not unique in general. But there is a unique optimal mapping which is monotone on the transfer rays [17].

In this paper we study the regularity of optimal mappings in Monge's mass transfer problem. As the cost function (1.1) does not satisfy condition **(B)**, the argument in [26] does not apply to Monge's problem. Indeed, Monge's problem also admits several minimizers  $T$ , even if a special one plays an important role: it is the transport map which is monotone on each transport ray (see [1]: we will call this map monotone optimal transport).

The regularity seems a rather tricky problem and very little is known at the moment. Only in the 2 dimensional case, it was proved in [19] that the monotone optimal mapping is continuous in the interior of the transfer set (i.e. the union of all transfer rays), under the assumptions that the densities  $f, g$  are positive, continuous, and with compact, convex and disjoint supports.

Our strategy to attack the regularity in Monge's problem is to establish uniform estimates for the optimal mappings with respect to the cost function

$$(1.3) \quad c_\varepsilon(x, y) = \sqrt{\varepsilon^2 + |x - y|^2}$$

where  $\varepsilon \in (0, 1]$  is a constant. The cost function  $c_\varepsilon$  satisfies both conditions **(A)** and **(B)**. Therefore there is a unique optimal mapping  $T_\varepsilon$  associated with  $c_\varepsilon$ , given by

$$T_\varepsilon(x) = x - \frac{\varepsilon Du_\varepsilon}{\sqrt{1 - |Du_\varepsilon|^2}}.$$

where  $u_\varepsilon$  is the potential function. By direct computation,  $u_\varepsilon$  satisfies the Monge-Ampère equation [26]

$$(1.4) \quad \det w_{ij} = \frac{1}{\varepsilon^n} \left[ 1 - |Du|^2 \right]^{\frac{n+2}{2}} \frac{f}{g},$$

with

$$\{w_{ij}\} = \left\{ \frac{\sqrt{1 - |Du|^2}}{\varepsilon} (\delta_{ij} - u_{x_i} u_{x_j}) - u_{x_i x_j} \right\}.$$

Under appropriate assumptions, the a priori estimate

$$(1.5) \quad \sup_{\Omega'} |D^2 u_\varepsilon(x)| \leq C_\varepsilon \quad \forall \Omega' \subset \subset \Omega.$$

was established in [26], where the upper bound  $C_\varepsilon$  depends on  $\varepsilon$ . Notice that the assumptions involve in particular lower bounds on the densities  $f$  and  $g$  on their respective domains  $\bar{\Omega}$  and  $\bar{\Omega}^*$ . these domains should be  $c_\varepsilon$ -convex w.r.t. each other, which typically reduces (if we want to impose it for all  $\varepsilon \rightarrow 0$ ) to the case of  $\Omega \subset \Omega^*$ , with  $\Omega^*$  convex. In particular, this rules out the assumption of [19], since the supports will not be disjoint. The case we study is thus completely different from that of [19].

Equation (1.4) is strongly singular as  $\varepsilon \rightarrow 0$ . Note that, due to the small  $\varepsilon$ , a uniform bound for  $D^2u_\varepsilon$  does not mean a uniform estimate for the optimal mapping  $T_\varepsilon$ . Therefore we need to work directly on the mapping  $T_\varepsilon$ .

We wished to prove a uniform bound for  $DT_\varepsilon$ , namely the uniform Lipschitz continuity of the optimal mapping  $T_\varepsilon$ . By tedious computations, we are able to prove that all the eigenvalues of the matrix  $DT_\varepsilon$ , which are all positive, are locally uniformly bounded as  $\varepsilon \rightarrow 0$ . This is one of the two main results of the paper. Notice that this should bring some information on the behavior of  $DT_0$ , where  $T_0$  is the monotone optimal mapping in Monge's problem. Yet, two problems arise: i) the condition on the eigenvalues being strongly nonlinear and applied to non-symmetric matrices, it is not easy to pass it to the limit, nor to give a meaning to the eigenvalues of  $DT_0$  (which is a priori a distribution); ii) even the fact that the maps  $T_\varepsilon$  do converge to the monotonic optimal transport is not that easy if the supports of the measures are not disjoint (which is not the case for us).

However, as the matrix  $DT_\varepsilon$  is - as we said - not symmetric, the boundedness of the eigenvalues of  $DT_\varepsilon$  does not imply the matrix itself is uniformly bounded. Interestingly, we find that the matrix  $DT_0$  is not bounded in general. There exist positive and smooth  $f, g$  such that  $DT_0$  is unbounded at interior points (here by  $T_0$  we mean the monotonic Monge optimal transport, and not the limit of  $T_\varepsilon$ ; however, it is possible to prove (see Section 4) that, should  $DT_\varepsilon$  be bounded, then  $T_\varepsilon \rightarrow T_0$ , and hence this implies that  $DT_\varepsilon$  cannot be uniformly bounded as  $\varepsilon \rightarrow 0$ ). This is the second main result of the paper.

This paper is arranged as follows. In section 2, we state our main estimate, Theorem 1. Section 3 is then devoted to the proof of Theorem 1. In section 4, we provide positive and smooth densities  $f, g$  such that the monotonic optimal mapping  $T_0$  is not Lipschitz continuous at interior points. We conclude the paper with some remarks and perspectives in Section 5.

## 2. UNIFORM A PRIORI ESTIMATES

Let  $c = c_\varepsilon$  be the cost function given in (1.3). The optimal mapping  $T = T_\varepsilon : \Omega \rightarrow \Omega^*$  is given by [26]

$$(2.1) \quad T(x) = [D_x c(x, \cdot)]^{-1} Du(x),$$

where  $u = u_\varepsilon$  is a  $c$ -concave potential function. In this and the next sections, we deal with the a priori estimates for  $DT$ . We will omit the subscript  $\varepsilon$  when no confusions arise.

From [26], the potential function  $u$  satisfies the fully nonlinear PDE of Monge-Ampère type,

$$(2.2) \quad \det(D_x^2 c - D^2 u) = |\det D_{xy}^2 c| \frac{f}{g \circ T} \text{ in } \Omega.$$

For the cost function (1.3), one has

$$(2.3) \quad Dc(x, y) = \frac{x - y}{\sqrt{\varepsilon^2 + |x - y|^2}}.$$

Hence by (2.1),

$$(2.4) \quad T(x) = x - L(x)Du(x),$$

where the function  $L$  is given by

$$(2.5) \quad L(x) =: \frac{\varepsilon}{\sqrt{1 - v}}$$

and

$$(2.6) \quad v =: |Du|^2.$$

From (2.4) and (2.5), we can solve

$$(2.7) \quad v = \frac{d^2(x)}{\varepsilon^2 + d^2(x)}$$

and consequently

$$(2.8) \quad L = \sqrt{\varepsilon^2 + d^2(x)},$$

where

$$(2.9) \quad d(x) = |x - T(x)|.$$

As in [26], we denote

$$(2.10) \quad \begin{aligned} A_{ij}(x) &= D_{x^i x^j}^2 c(x, T(x)) \\ &= \frac{1}{L}(\delta_{ij} - D_i u D_j u). \end{aligned}$$

Then equation (2.2) can be written in the form

$$(2.11) \quad \det w_{ij} = \frac{\varepsilon^2}{L^{n+2}} \frac{f}{g \circ T},$$

where

$$(2.12) \quad w_{ij} =: A_{ij} - D_{ij}^2 u$$

is a nonnegative symmetric matrix.

We observe from (2.10) that  $A_{ij}$  is positive definite, and the inverse matrix of  $A_{ij}$  is given by

$$(2.13) \quad A^{ij} = L\left(\delta_{ij} + \frac{L^2}{\varepsilon^2} D_i u D_j u\right).$$

Let us denote

$$(2.14) \quad W =: \sum_{\alpha, \beta=1}^n A^{\alpha\beta} w_{\alpha\beta}.$$

Then we have following uniform estimates:

**Theorem 1.** *Suppose  $\Omega, \Omega^*$  are bounded domains in  $\mathbb{R}^n$  ( $n \geq 2$ ),  $f \in C^{1,1}(\Omega)$ ,  $g \in C^{1,1}(\Omega^*)$ ,  $f, g$  have positive upper and lower bounds, and (1.2) holds. Let  $u \in C^{3,1}(\Omega)$  be a  $c$ -concave solution to (2.11), then we have a priori estimate*

$$(2.15) \quad W(x) \leq C,$$

where  $C$  depends on  $n$ ,  $\text{dist}(x, \partial\Omega)$ ,  $f$  and  $g$ , but is independent of the constant  $\varepsilon \in (0, 1]$ .

By (2.4) (2.12) and (2.13), it is ready to check that the Jacobian matrix of  $T$  is given by

$$(2.16) \quad \begin{aligned} T_j^i &= \delta_{ij} - L_j u_i - L u_{ij} \\ &= \delta_{ij} - L(u_{ij} + \frac{L^2}{\varepsilon^2} u_i \sum_k u_k u_{kj}) \\ &= \sum_k A^{ik} w_{kj}. \end{aligned}$$

Since the matrices  $\{A^{ij}\}$  and  $\{w_{ij}\}$  are positive, then  $DT$  is diagonalizable, and its eigenvalues  $\lambda_1, \dots, \lambda_n$  of Jacobian  $DT$  are positive, and  $\sum_{i=1}^n \lambda_i = W$ . So if  $W$  is bounded, one immediately sees that all the eigenvalues of  $DT$  are bounded from above and below. We therefore have

**Corollary 1.** *Under the assumptions of Theorem 1, we have for any  $\Omega' \subset \subset \Omega$ ,*

$$(2.17) \quad C^{-1} \leq \min_i \lambda_i \leq \max_i \lambda_i \leq C \quad \text{in } \Omega',$$

where  $C$  depends on  $n$ ,  $\text{dist}(\Omega', \partial\Omega)$ ,  $f$  and  $g$ , but is independent of  $\varepsilon \in (0, 1]$ .

In view of (2.13) and (2.14), one finds that

$$W = L \sum_i w_{ii} + \frac{L^3}{\varepsilon^2} \sum_{i,j} u_i u_j w_{ij}.$$

Hence we obtain

$$L \sum w_{ii} \leq C$$

By (2.10) and (2.12), we also obtain the estimate for  $D^2u$ .

**Corollary 2.** *Under the assumptions of Theorem 1, we have for any  $\Omega' \subset\subset \Omega$ ,*

$$(2.18) \quad |D^2u| \leq C/L \quad \text{in } \Omega',$$

where  $C$  depends on  $n$ ,  $\text{dist}(\Omega', \partial\Omega)$ ,  $f$  and  $g$ , but is independent of  $\varepsilon \in (0, 1]$ .

Corollary 2 morally gives a  $C^{1,1}$  estimate for the potential function  $u_0 = \lim_{\varepsilon \rightarrow 0} u$  on the set

$$E_\delta = \{x \in \Omega : |T(x) - x| \geq \delta > 0\}.$$

This recovers a well-known result which reads “the potential is  $C^{1,1}$  in the interior of transport rays”, which was also used by [1] in order to prove the countable Lipschitz property of the direction of  $Du$ . At a point  $x_0$  with  $v(x_0) > 0$ , denote

$$\nu = -\frac{Du(x_0)}{|Du(x_0)|},$$

and let  $\xi^\alpha$  be unit vectors such that  $\{\nu, \xi^\alpha\}_{\alpha=1, \dots, n-1}$  are orthonormal. We denote

$$\begin{aligned} T_\nu^\nu &= \sum_{i,j} \nu_i \nu_j T_j^i, \\ T_{\xi^\alpha}^{\xi^\alpha} &= \sum_{i,j} \xi_i^\alpha \xi_j^\alpha T_j^i. \end{aligned}$$

By (2.16) and (2.13),

$$\begin{aligned} (2.19) \quad D_\nu \langle \nu, T \rangle &= T_\nu^\nu = \sum_{i,j,k} \nu_i A^{ik} w_{kj} \nu_j \\ &= L \sum_{i,j,k} \left(1 + \frac{vL^2}{\varepsilon^2}\right) \nu_k w_{kj} \nu_j \\ &= \frac{L^3}{\varepsilon^2} \sum_{j,k} \nu_k w_{kj} \nu_j. \end{aligned}$$

Similarly,

$$(2.20) \quad D_{\xi^\alpha} \langle \xi^\alpha, T \rangle = T_{\xi^\alpha}^{\xi^\alpha} = L \sum_{j,k} \xi_k^\alpha w_{kj} \xi_j^\alpha.$$

Noticing that  $\{w_{ij}\}$  is positive definite, it is clear from (2.19) and (2.20) that  $T_\nu^\nu$  and  $T_{\xi^\alpha}^{\xi^\alpha}$  are positive. Recall that

$$W = T_\nu^\nu + \sum_{\alpha=1}^{n-1} T_{\xi^\alpha}^{\xi^\alpha}.$$

Hence by (2.15) we obtain

**Corollary 3.** *Under the assumptions of Theorem 1, we have for any  $\Omega' \subset\subset \Omega$ ,*

$$(2.21) \quad \begin{cases} T_\nu^\nu \leq C \\ T_{\xi^\alpha}^{\xi^\alpha} \leq C \end{cases} \quad \text{in } \Omega',$$

where  $C$  depends on  $n$ ,  $\text{dist}(\Omega', \partial\Omega)$ ,  $f$  and  $g$ , but is independent of  $\varepsilon \in (0, 1]$ .

At the limit, this corresponds to saying (even if what we state here is not rigorous) that the limit mapping  $T_0$  is Lipschitz continuous in the direction of transfer rays, and for any unit vector  $\xi$  perpendicular to the transfer rays,  $\langle \xi, T_0 \rangle$  is Lipschitz continuous in the  $\xi$ -direction. The Lipschitz continuity along transport rays is not surprising, since we are doing one-dimensional optimal transport between two measures with upper and lower bounds on their densities; yet, the densities of the one-dimensional problem along each ray are affected by a Jacobian factor (due to the decomposition of  $f$  and  $g$  along rays), and this makes this Lipschitz result not completely evident. In section 4, we will construct an example to show that the component  $\langle \nu_0, T_0 \rangle$  is in general not Lipschitz continuous in  $\xi$ , even though the mass distributions are positive and smooth, where  $\nu_0$  is a direction of transfer rays and  $\xi \perp \nu_0$ .

In Theorem 1, we assume that  $u \in C^{3,1}$ . This assumption is not needed if  $\Omega \subset \Omega^*$  and  $\Omega^*$  is a bounded convex domain in  $\mathbb{R}^n$ , as it implies that  $\Omega^*$  is  $c^*$ -convex with respect to  $\Omega$  and by approximation, and the condition  $u \in C^{3,1}$  is always satisfied, see [26].

### 3. PROOF OF THEOREM 1

To prove Theorem 1, we introduce the auxiliary function

$$(3.1) \quad H(x) = \eta(x)W(x),$$

where  $\eta$  is a cut-off function. Suppose that  $H$  attains its maximum at some point  $x_0$ . To prove that  $H(x_0)$  is uniformly bounded in  $\varepsilon$ , the computation is rather complicated. We find the computation can be made a little simpler if we first make a linear transformation such that

$$(3.2) \quad A_{ij}(x_0) = \delta_{ij},$$

and then make a rotation of coordinates such that

$$(3.3) \quad w_{ij}(x_0) = \text{diag} \{ \lambda_1, \dots, \lambda_n \}.$$



It is well-known that  $A_{ij}$ ,  $w_{ij}$  are tensors [22]. An advantage of working on tensors is that one may choose a particular coordinate system to simplify the computation. As we only made a linear transform on the Euclidean space  $\mathbb{R}^n$ , the Riemannian curvature tensor under the metric  $\sigma_{ij}$  vanishes, which allows us to exchange the derivatives freely. In the following we will use  $D$  to denote the normal derivatives in  $\mathbb{R}^n$  and  $\nabla$  to denote covariant derivatives under the metric  $\sigma$ .

Suppose the linear transformation is given by  $y = P^{-1}x$  (i.e.  $x_i = \sum P_{ik}y_k$ ) such that  $P^TAP = I$  is the unit matrix at  $x_0$ . Then by (2.10) and (2.12),

$$\begin{aligned}\bar{A}_{ij} &= \sum_{k,l} A_{kl}P_{ki}P_{lj} = (P^TAP)_{ij}, \\ \bar{w}_{ij} &= \sum_{k,l} w_{kl}P_{ki}P_{lj} = (P^TwP)_{ij},\end{aligned}$$

where bar denotes quantities in the  $y$ -coordinates.

Denote  $\{\sigma_{ij}\} = P^TP$ , and  $\{\sigma^{ij}\} = (P^TP)^{-1}$ . Then by (2.10) and (2.13),

$$\begin{aligned}(3.4) \quad \delta_{ij} &= \bar{A}_{ij} = (P^TAP)_{ij} \\ &= \frac{1}{L}(\sigma_{ij} - \bar{u}_i\bar{u}_j)\end{aligned}$$

and

$$\begin{aligned}(3.5) \quad \delta_{ij} &= \bar{A}^{ij} = (P^{-1}A^{-1}(P^T)^{-1})_{ij} \\ &= L(\sigma^{ij} + \frac{L^2}{\varepsilon^2}\bar{u}^i\bar{u}^j),\end{aligned}$$

where  $\bar{u}^i =: \sum_k \sigma^{ik}\bar{u}_k$ . Note that by (2.7) and (2.8),  $v$  and  $L$  are invariant under the coordinate transformation, and

$$(3.6) \quad \bar{v} = \sum \sigma^{ij}\bar{u}_i\bar{u}_j = \sum \bar{u}^i\bar{u}_i \leq 1,$$

$$(3.7) \quad \varepsilon^2 \leq \bar{L}^2 \leq C.$$

For simplicity we will omit the bar below. In view of (3.4), we have, at  $x_0$ ,

$$(3.8) \quad \sigma_{ij} = L\delta_{ij} + \nabla_i u \nabla_j u.$$

By (3.5), we have, at  $x_0$ ,

$$\begin{aligned}u_i &= \sum_j \delta_{ij}u_j \\ &= \sum_j L(\sigma^{ij} + \frac{L^2}{\varepsilon^2}u^i u^j)u_j \\ &= L(1 + \frac{L^2}{\varepsilon^2}v)u^i,\end{aligned}$$

where  $u_i = \nabla_i u$ . By (2.5), it follows that

$$(3.9) \quad u_i = \frac{Lu^i}{1-v} = \frac{L^3}{\varepsilon^2} u^i.$$

Hence  $u^i = \frac{\varepsilon^2}{L^3} u_i$  and by (3.5) ,

$$(3.10) \quad \sigma^{ij} = \frac{1}{L} \left( \delta_{ij} - \frac{\varepsilon^2}{L^3} u_i u_j \right).$$

Formulas (3.8), (3.9) and (3.10) will be repeatedly used in our calculation below. Without loss of generality, we may also assume

$$(3.11) \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

at  $x_0$ .

Since  $x_0$  is the maximum point, we have

$$(3.12) \quad \begin{aligned} 0 &= \nabla_i \log H(x_0) = \frac{\eta_i}{\eta} + \frac{W_i}{W} \\ &= \frac{\eta_i}{\eta} + \frac{\sum w_{\alpha\alpha;i}}{W} - \frac{\sum A_{\alpha\alpha;i} w_{\alpha\alpha}}{W}, \\ 0 &\geq \nabla_{ij}^2 \log H(x_0) = \frac{\eta_{ij}}{\eta} - 2 \frac{\eta_i \eta_j}{\eta^2} + \frac{W_{ij}}{W} \end{aligned}$$

as a matrix, where subscripts  $i, j$  on the R.H.S. denote covariant derivatives in the metric  $\sigma$ . We thus obtain, at  $x_0$ ,

$$(3.13) \quad 0 \geq W \sum w^{ij} \left( \frac{\eta_{ij}}{\eta} - 2 \frac{\eta_i \eta_j}{\eta^2} \right) + \sum w^{ij} W_{ij},$$

where  $w^{ij}$  is the inverse of  $w_{ij}$ .

Differentiating (2.11) gives

$$(3.14) \quad \sum w^{ij} w_{ij;a} = \varphi_a,$$

$$(3.15) \quad \sum w^{ij} w_{ij;ab} = \sum w^{is} w^{jt} w_{ij;a} w_{st;b} + \varphi_{ab},$$

where  $\varphi$  is given by

$$(3.16) \quad \varphi = \log \left( \frac{\varepsilon^2}{L^{n+2}} \frac{f}{g \circ T} \right).$$

In our computation we use the notation  $w_{ij;k} = \nabla_k w_{ij}$ ,  $w_{ij;kl} = \nabla_l \nabla_k w_{ij}$ ,  $A_{ij;k} = \nabla_k A_{ij}$  and  $A_{ij;kl} = \nabla_l \nabla_k A_{ij}$ .

To estimate the term  $\sum w^{ij} W_{ij}$  in (3.13), we first prove the following lemma.

**Lemma 1.** *We have*

$$\begin{aligned} A_{ij;k} &= \frac{L^2}{\varepsilon^2} A_{ij} u^h w_{hk} + \frac{1}{L} (u_j w_{ik} + u_i w_{jk}) \\ &\quad - \frac{1}{L} (A_{ij} u_k + A_{ik} u_j + u_i A_{jk}), \\ A_{ii;\beta} - A_{i\beta;i} &= \frac{L^2}{\varepsilon^2} (A_{ii} u^t w_{t\beta} - A_{i\beta} u^t w_{ti}) + \frac{1}{L} (w_{i\beta} u_i - w_{ii} u_\beta), \end{aligned}$$

(we use the summation convention  $u^h w_{hk} = \sum_h u^h w_{hk}$ ).

*Proof.* Recall that  $v = \sigma^{ij} u_i u_j = \sum u^i u_i$ . Therefore

$$\begin{aligned} \frac{dL}{dv} &= \frac{1}{2} \frac{L^3}{\varepsilon^2}, \\ v_k &= 2u^h u_{hk}. \end{aligned}$$

By (3.4),  $A_{ij} = \frac{1}{L} (\sigma_{ij} - u_i u_j)$ . Differentiating, we get

$$\begin{aligned} (3.17) \quad A_{ij;k} &= -\frac{1}{L^2} \frac{dL}{dv} v_k (\sigma_{ij} - u_i u_j) + \frac{1}{L} (-u_{ik} u_j - u_i u_{jk}) \\ &= -\frac{L^2}{\varepsilon^2} A_{ij} u^h u_{hk} + \frac{1}{L} (-u_{ik} u_j - u_i u_{jk}) \\ &= \frac{L^2}{\varepsilon^2} A_{ij} u^h w_{hk} + \frac{1}{L} (w_{ik} u_j + u_i w_{jk}) \\ &\quad - \frac{1}{L} (A_{ij} u_k + A_{ik} u_j + u_i A_{jk}). \end{aligned}$$

The second formula follows from (3.17) immediately.  $\square$

Differentiating (2.14) twice and using  $A_{ij}(x_0) = \delta_{ij}$ ,

$$\begin{aligned} (3.18) \quad \sum_{i,j} w^{ij} W_{ij} &= \sum w^{ij} w_{\alpha\alpha;ij} - 2 \sum w^{ij} A_{\alpha\beta;i} w_{\alpha\beta;j} - \sum w^{ij} A_{\alpha\alpha;ij} w_{\alpha\alpha} \\ &\quad + 2 \sum w^{ij} A_{\alpha k;i} A_{\beta k;j} w_{\alpha\beta} \\ &\geq \sum w^{ij} w_{\alpha\alpha;ij} - 2 \sum w^{ij} A_{\alpha\beta;i} w_{\alpha\beta;j} - \sum w^{ij} A_{\alpha\alpha;ij} w_{\alpha\alpha}. \end{aligned}$$

We have by (3.15)

$$\begin{aligned} (3.19) \quad \sum_{i,j,\alpha} w^{ij} w_{\alpha\alpha;ij} &= \sum w^{ij} A_{\alpha\alpha;ij} - \sum w^{ij} u_{\alpha\alpha ij} \\ &= \sum w^{ij} w_{ij;\alpha\alpha} + \sum w^{ij} (A_{\alpha\alpha;ij} - A_{ij;\alpha\alpha}) \\ &\geq \sum \varphi_{\alpha\alpha} + \sum w^{ii} (A_{\alpha\alpha;ii} - A_{ii;\alpha\alpha}). \end{aligned}$$

By the first formula in Lemma 1

$$\begin{aligned} \sum_{i,j,\alpha,\beta} w^{ij} A_{\alpha\beta;i} w_{\alpha\beta;j} &= \frac{L^2}{\varepsilon^2} \sum w_{\alpha\alpha;i} u^i - \frac{1}{L} \sum w^{ij} u_j w_{\alpha\alpha;i} \\ &\quad + \frac{2}{L} \sum u_\beta w_{\alpha\beta;\alpha} - \frac{2}{L} \sum w^{ij} u_\beta w_{j\beta;i}. \end{aligned}$$

By (3.9), it follows

$$\begin{aligned} \sum_{i,j,\alpha,\beta} w^{ij} A_{\alpha\beta;i} w_{\alpha\beta;j} &= 3 \frac{L^2}{\varepsilon^2} \sum w_{\alpha\alpha;i} u^i - \frac{1}{L} \sum w^{ij} u_j w_{\alpha\alpha;i} \\ &\quad + \frac{2}{L} \sum u_\beta (A_{\alpha\beta;\alpha} - A_{\alpha\alpha;\beta}) - \frac{2}{L} \sum w^{ij} u_\beta w_{ij;\beta} \\ &\quad - \frac{2}{L} \sum w^{ii} u_\beta (A_{i\beta;i} - A_{ii;\beta}). \end{aligned}$$

By (3.14) and the second formula in Lemma 1, we then obtain

$$\begin{aligned} \sum_{i,j,\alpha,\beta} w^{ij} A_{\alpha\beta;j} w_{\alpha\beta;i} &= 3 \frac{L^2}{\varepsilon^2} \sum w_{\alpha\alpha;i} u^i - \frac{1}{L} \sum w^{ii} u_i w_{\alpha\alpha;i} \\ &\quad - 2 \frac{L^2}{\varepsilon^2} \sum \varphi_\beta u^\beta + 2 \frac{Lv}{\varepsilon^2} (W - n) \\ &\quad + 2 \frac{L}{\varepsilon^2} (\mathcal{W} - n) \sum w_{ii} u_i u^i, \end{aligned}$$

where

$$\mathcal{W} =: \sum w^{ii} = \sum \frac{1}{\lambda_i}.$$

Recalling (3.9) and (3.6),

$$(3.20) \quad 0 \leq u^i u_i \leq \sum u^i u_i \leq 1$$

for any given  $i$ . Hence

$$\begin{aligned} (3.21) \quad \sum_{i,j,\alpha,\beta} w^{ij} A_{\alpha\beta;j} w_{\alpha\beta;i} &\leq 3 \frac{L^2}{\varepsilon^2} \sum w_{\alpha\alpha;i} u^i - \frac{1}{L} \sum w^{ii} u_i w_{\alpha\alpha;i} \\ &\quad - 2 \frac{L^2}{\varepsilon^2} \sum \varphi_\beta u^\beta + \frac{L}{\varepsilon^2} Q. \end{aligned}$$

Here and below we use  $Q$  to denote quantities satisfying

$$Q \leq C \left( 1 + \frac{W}{\eta} + W^2 + \frac{1}{\eta} W \mathcal{W} \right).$$

Inserting (3.19) and (3.21) into (3.18), we obtain

$$(3.22) \quad \begin{aligned} \sum w^{ij} W_{ij} &\geq \sum w^{ii} (A_{\alpha\alpha;ii} - A_{ii;\alpha\alpha}) - \sum w^{ii} A_{\alpha\alpha;ii} w_{\alpha\alpha} \\ &\quad - 6 \frac{L^2}{\varepsilon^2} \sum w_{\alpha\alpha;i} u^i + \frac{2}{L} \sum w^{ii} u_i w_{\alpha\alpha;i} \\ &\quad + \sum \varphi_{\alpha\alpha} + \frac{4L^2}{\varepsilon^2} \sum \varphi_{\alpha} u^{\alpha} - \frac{L}{\varepsilon^2} Q. \end{aligned}$$

To proceed further, we need the following lemma.

**Lemma 2.** *We have*

$$\begin{aligned} \sum_{i,\alpha} w^{ii} (A_{\alpha\alpha;ii} - A_{ii;\alpha\alpha}) &\geq \frac{L}{\varepsilon^2} \mathcal{W} \sum w_{ii}^2 - \frac{2}{L} \sum w^{ii} u_i w_{\alpha\alpha;i} \\ &\quad - \frac{L^2}{\varepsilon^2} \mathcal{W} \sum w_{\alpha\alpha;h} u^h + (n+2) \frac{L^2}{\varepsilon^2} \sum \varphi_{\beta} u^{\beta} - \frac{L}{\varepsilon^2} Q \end{aligned}$$

and

$$\begin{aligned} \sum_{i,\alpha} w^{ii} A_{\alpha\alpha;ii} w_{\alpha\alpha} &\leq -2 \frac{L}{\varepsilon^2} \mathcal{W} \sum w_{ii}^2 u_i u^i - \frac{L}{\varepsilon^2} \mathcal{W} W \sum w_{ii} u_i u^i \\ &\quad + \frac{L^2}{\varepsilon^2} W \sum \varphi_{\beta} u^{\beta} + \frac{2L^2}{\varepsilon^2} \sum w_{ii} u^i \varphi_i + \frac{L}{\varepsilon^2} Q. \end{aligned}$$

*Proof.* In view of Lemma 1,

$$(3.23) \quad A_{\alpha\alpha;i} = -\frac{L^2}{\varepsilon^2} A_{\alpha\alpha} u^h u_{hi} - \frac{2}{L} u_{\alpha} u_{\alpha i}.$$

By differentiating (3.23),

$$(3.24) \quad \begin{aligned} A_{\alpha\alpha;ii} &= -\frac{L^2}{\varepsilon^2} A_{\alpha\alpha} u^h u_{hii} - \frac{L^2}{\varepsilon^2} A_{\alpha\alpha} \sigma^{ht} u_{ti} u_{hi} \\ &\quad - \frac{L^2}{\varepsilon^2} A_{\alpha\alpha;i} u^h u_{hi} - 2 \frac{L^4}{\varepsilon^4} A_{\alpha\alpha} u^t u_{ti} u^h u_{hi} \\ &\quad - \frac{2}{L} u_{\alpha} u_{i\alpha i} - \frac{2}{L} u_{i\alpha}^2 + \frac{2L}{\varepsilon^2} u^h u_{hi} u_{\alpha} u_{i\alpha}. \end{aligned}$$

Plugging (3.23) into (3.24), we infer that

$$\begin{aligned} A_{\alpha\alpha;ii} &= -\frac{L^2}{\varepsilon^2} A_{\alpha\alpha} u^h u_{iih} - \frac{L^2}{\varepsilon^2} A_{\alpha\alpha} \sigma^{ht} u_{ti} u_{hi} \\ &\quad - \frac{L^4}{\varepsilon^4} A_{\alpha\alpha} u^t u_{ti} u^h u_{hi} - \frac{2}{L} u_{\alpha} u_{ii\alpha} \\ &\quad - \frac{2}{L} u_{i\alpha}^2 + \frac{4L}{\varepsilon^2} u^h u_{hi} u_{\alpha} u_{i\alpha}. \end{aligned}$$

By (3.10),

$$-\frac{L^2}{\varepsilon^2} A_{\alpha\alpha} \sigma^{ht} u_{ti} u_{hi} = -\frac{L}{\varepsilon^2} A_{\alpha\alpha} u_{ii}^2 + \frac{L^4}{\varepsilon^4} A_{\alpha\alpha} u_{ii}^2 u^i u^i.$$

Hence

$$\begin{aligned}
A_{\alpha\alpha;ii} &= -\frac{L^2}{\varepsilon^2}A_{\alpha\alpha}u^hu_{iih} - \frac{L}{\varepsilon^2}A_{\alpha\alpha}u_{ii}^2 \\
&\quad - \frac{2}{L}u_\alpha u_{ii\alpha} - \frac{2}{L}u_{i\alpha}^2 + \frac{4L}{\varepsilon^2}u_{ii}u_{i\alpha}u_\alpha u^i \\
&= \frac{L^2}{\varepsilon^2}A_{\alpha\alpha}u^hw_{ii;h} - \frac{L^2}{\varepsilon^2}A_{\alpha\alpha}u^hA_{ii;h} - \frac{L}{\varepsilon^2}A_{\alpha\alpha}u_{ii}^2 \\
&\quad + \frac{2}{L}u_\alpha w_{ii;\alpha} - \frac{2}{L}u_\alpha A_{ii;\alpha} - \frac{2}{L}u_{i\alpha}^2 + \frac{4L}{\varepsilon^2}u_{ii}u_{i\alpha}u_\alpha u^i.
\end{aligned}$$

Employing (3.23) again, it follows

$$\begin{aligned}
(3.25) \quad A_{\alpha\alpha;ii} &= \frac{L^2}{\varepsilon^2}A_{\alpha\alpha}w_{ii;h}u^h + \frac{2}{L}u_\alpha w_{ii;\alpha} \\
&\quad - \frac{L}{\varepsilon^2}A_{\alpha\alpha}u_{ii}^2 - \frac{2}{L}u_{i\alpha}^2 + \frac{4L}{\varepsilon^2}u_{ii}u_{i\alpha}u_\alpha u^i \\
&\quad + \frac{L^4}{\varepsilon^4}A_{\alpha\alpha}A_{ii}u_{ht}u^hu^t + 2\frac{L}{\varepsilon^2}A_{\alpha\alpha}u_{ii}u_i u^i \\
&\quad + 2\frac{L}{\varepsilon^2}A_{ii}u_\alpha u_\alpha u^\alpha + \frac{4}{L^2}u_{i\alpha}u_\alpha u_i.
\end{aligned}$$

Hence

$$\begin{aligned}
(3.26) \quad \sum_{i,\alpha} w^{ii} A_{\alpha\alpha;ii} w_{\alpha\alpha} &= \frac{L^2}{\varepsilon^2}W \sum \varphi_h u^h + 2\frac{L^2}{\varepsilon^2} \sum w_{ii} u^i \varphi_i \\
&\quad - \frac{L}{\varepsilon^2}W \sum w^{ii} u_{ii}^2 - \frac{2}{L} \sum u_{ii}^2 \\
&\quad + \frac{L}{\varepsilon^2} \sum \left\{ 4u_{ii}^2 u_i u^i + 4u_{ii} u_i u^i + 2W w^{ii} u_{ii} u_i u^i \right\} \\
&\quad + \frac{L}{\varepsilon^2} \left\{ W\mathcal{W} \sum u_{ii} u_i u^i + 2\mathcal{W} \sum w_{ii} u_{ii} u_i u^i \right\}.
\end{aligned}$$

Recalling (3.20),

$$\sum \{4u_{ii}^2 u_i u^i + 4u_{ii} u_i u^i + 2W w^{ii} u_{ii} u_i u^i\} \leq Q,$$

and

$$W\mathcal{W} \sum u_{ii} u_i u^i + 2\mathcal{W} \sum w_{ii} u_{ii} u_i u^i \leq - \left\{ W\mathcal{W} \sum w_{ii} u_i u^i + 2\mathcal{W} \sum w_{ii}^2 u_i u^i \right\} + Q,$$

the second inequality of Lemma 2 follows from (3.26).

From (3.25) it follows that

$$\begin{aligned}
A_{\alpha\alpha;ii} - A_{ii;\alpha\alpha} &= \frac{L^2}{\varepsilon^2} A_{\alpha\alpha} w_{ii;h} u^h + \frac{2}{L} u_\alpha w_{ii;\alpha} \\
&\quad - \frac{L^2}{\varepsilon^2} A_{ii} w_{\alpha\alpha;h} u^h - \frac{2}{L} u_i w_{\alpha\alpha;i} \\
&\quad - \frac{L}{\varepsilon^2} A_{\alpha\alpha} u_{ii}^2 + \frac{4L}{\varepsilon^2} u_{ii} u_{i\alpha} u_\alpha u^i \\
&\quad + \frac{L}{\varepsilon^2} A_{ii} u_{\alpha\alpha}^2 - \frac{4L}{\varepsilon^2} u_{\alpha\alpha} u_{i\alpha} u_i u^\alpha.
\end{aligned}$$

Hence

$$\begin{aligned}
(3.27) \quad \sum w^{ii} (A_{\alpha\alpha;ii} - A_{ii;\alpha\alpha}) &= \frac{(n+2)L^2}{\varepsilon^2} \sum \varphi_i u^i \\
&\quad - \frac{L^2}{\varepsilon^2} \mathcal{W} \sum w_{\alpha\alpha;h} u^h - \frac{2}{L} \sum w^{ii} u_i w_{\alpha\alpha;i} \\
&\quad - \frac{nL}{\varepsilon^2} \sum w^{ii} u_{ii}^2 + \frac{L}{\varepsilon^2} \mathcal{W} \sum u_{ii}^2.
\end{aligned}$$

Since

$$-\frac{nL}{\varepsilon^2} \sum w^{ii} u_{ii}^2 \geq -\frac{L}{\varepsilon^2} Q$$

and

$$\frac{L}{\varepsilon^2} \mathcal{W} \sum u_{ii}^2 \geq \frac{L}{\varepsilon^2} \mathcal{W} \sum w_{ii}^2 - \frac{L}{\varepsilon^2} Q,$$

the first inequality of Lemma 2 follows from (3.27).  $\square$

In view of Lemma 2, (3.22) can be rewritten in the form

$$\begin{aligned}
(3.28) \quad \sum w^{ij} W_{ij} &\geq \frac{L}{\varepsilon^2} \mathcal{W} \sum w_{ii}^2 + \frac{L}{\varepsilon^2} \mathcal{W} (W \sum w_{ii} u_i u^i + 2 \sum w_{ii}^2 u_i u^i) \\
&\quad - \frac{L^2}{\varepsilon^2} (\mathcal{W} + 6) \sum w_{\alpha\alpha;i} u^i + \mathfrak{R}_\varphi - \frac{L}{\varepsilon^2} Q,
\end{aligned}$$

where

$$(3.29) \quad \mathfrak{R}_\varphi =: -\frac{2L^2}{\varepsilon^2} \sum w_{ii} u^i \varphi_i + (n+6-W) \frac{L^2}{\varepsilon^2} \sum \varphi_\beta u^\beta + \sum \varphi_{\alpha\alpha}.$$

By (3.12), we have

$$(3.30) \quad \sum_\alpha w_{\alpha\alpha;k} = \sum A_{\alpha\alpha;k} w_{\alpha\alpha} - W \frac{\eta_k}{\eta}.$$

It follows from (3.23) that

$$\begin{aligned}
 (3.31) \quad \sum_{\alpha} w_{\alpha\alpha;k} &= -W \frac{\eta_k}{\eta} - \frac{L^2}{\varepsilon^2} W u_{kk} u^k - \frac{2}{L} w_{kk} u_{kk} u_k \\
 &= -W \frac{\eta_k}{\eta} + \frac{L^2}{\varepsilon^2} W w_{kk} u^k - \frac{L^2}{\varepsilon^2} W u^k \\
 &\quad + \frac{2L^2}{\varepsilon^2} w_{kk}^2 u^k - \frac{2L^2}{\varepsilon^2} w_{kk} u^k.
 \end{aligned}$$

Hence, by (3.9) and (3.20)

$$\begin{aligned}
 (3.32) \quad -\frac{L^2}{\varepsilon^2} (\mathcal{W} + 6) \sum w_{\alpha\alpha;i} u^i &\geq -\frac{L}{\varepsilon^2} Q + \frac{L^2}{\varepsilon^2} W (\mathcal{W} + 6) \frac{\sum u^i \eta_i}{\eta} \\
 &\quad - \frac{L}{\varepsilon^2} \mathcal{W} (W \sum w_{ii} u^i u_i + 2 \sum w_{ii}^2 u_i u^i).
 \end{aligned}$$

Therefore, by inserting (3.32) into (3.28), we find that (3.13) can be written as

$$\begin{aligned}
 (3.33) \quad 0 &\geq \frac{L}{\varepsilon^2} \mathcal{W} \sum w_{ii}^2 + W \sum w^{ij} \left( \frac{\eta_{ij}}{\eta} - 2 \frac{\eta_i \eta_j}{\eta^2} \right) \\
 &\quad + \frac{L^2}{\varepsilon^2} W (\mathcal{W} + 6) \frac{\sum u^i \eta_i}{\eta} + \mathfrak{R}_{\varphi} - \frac{L}{\varepsilon^2} Q.
 \end{aligned}$$

Without loss of generality, we may assume the cut-off function  $\eta$  satisfies  $|D\eta|^2 \leq C\eta$  (otherwise we may replace  $\eta$  by  $\eta^2$ ) and  $|D^2\eta| \leq C$ . Hence it follows

$$\begin{aligned}
 &W \sum w^{ij} \left( \frac{\eta_{ij}}{\eta} - 2 \frac{\eta_i \eta_j}{\eta^2} \right) + \frac{L^2}{\varepsilon^2} W (\mathcal{W} + 6) \frac{\sum u^i \eta_i}{\eta} \\
 &\geq -C \frac{W}{\eta} (|D^2\eta| + \frac{|D\eta|^2}{\eta}) \sum w^{ij} \sigma_{ij} - C \frac{L^2}{\varepsilon^2} \frac{W}{\eta} (\mathcal{W} + 1) \\
 &\geq -\frac{L}{\varepsilon^2} Q,
 \end{aligned}$$

where (3.8) is used in the last inequality. Therefore (3.33) can be written as

$$(3.34) \quad 0 \geq \frac{L}{\varepsilon^2} \mathcal{W} \sum w_{ii}^2 + \mathfrak{R}_{\varphi} - \frac{L}{\varepsilon^2} Q.$$

**Lemma 3.** *We have, at  $x_0$ ,*

$$\mathfrak{R}_{\varphi} \geq -\frac{L}{\varepsilon^2} Q.$$

*Proof.* Recalling (2.6), we have

$$(3.35) \quad v_{\alpha} = 2u^h u_{h\alpha}.$$



Differentiating (3.16) gives

$$(3.36) \quad \varphi_\alpha = -(n+2) \frac{L^2}{\varepsilon^2} u^h u_{h\alpha} + \frac{f_\alpha}{f} - \frac{\nabla_\alpha(g \circ T)}{g \circ T},$$

$$(3.37) \quad \begin{aligned} \varphi_{\alpha\beta} = & -2(n+2) \frac{L^4}{\varepsilon^4} u^h u^t u_{h\alpha} u_{t\beta} - \frac{(n+2)}{2} \frac{L^2}{\varepsilon^2} v_{\alpha\beta} + \frac{f_{\alpha\beta}}{f} \\ & - \frac{f_\alpha f_\beta}{f^2} - \frac{\nabla_{\alpha\beta}^2(g \circ T)}{g \circ T} + \frac{\nabla_\alpha(g \circ T)}{g \circ T} \frac{\nabla_\beta(g \circ T)}{g \circ T}. \end{aligned}$$

Inserting (3.36) and (3.37) into (3.29), we obtain

$$(3.38) \quad \begin{aligned} \mathfrak{R}_\varphi \geq & 2 \frac{L^2}{\varepsilon^2} \sum w_{ii} \frac{\nabla_i(g \circ T)}{g \circ T} u^i - (n+6-W) \frac{L^2}{\varepsilon^2} \sum \frac{u^\alpha \nabla_\alpha(g \circ T)}{g \circ T} \\ & - \sum \frac{\nabla_{\alpha\alpha}^2(g \circ T)}{g \circ T} - \frac{(n+2)}{2} \frac{L^2}{\varepsilon^2} \sum v_{\alpha\alpha} - \frac{L}{\varepsilon^2} Q. \end{aligned}$$

Differentiating (3.35), we obtain

$$\begin{aligned} \sum v_{\alpha\alpha} &= 2 \sum \sigma^{th} u_{t\alpha} u_{h\alpha} + 2 \sum u^h u_{\alpha\alpha h} \\ &= 2 \sum \sigma^{th} u_{t\alpha} u_{h\alpha} + 2 \sum u^h A_{\alpha\alpha;h} - 2 \sum u^h w_{\alpha\alpha;h}. \end{aligned}$$

By (3.9), (3.10) and (3.20),

$$\sum \sigma^{th} u_{t\alpha} u_{h\alpha} = \frac{1}{L} \sum (u_{ii}^2 - u_{ii}^2 u^i u_i) \leq \frac{1}{L} Q.$$

From (3.23),

$$\sum u^i A_{\alpha\alpha;i} = -\frac{n+2}{L} u_{ii} u^i u_i \leq \frac{1}{L} Q.$$

Also, by (3.31),

$$\begin{aligned} -\sum u^k w_{\alpha\alpha;k} &= W \sum \frac{u^k \eta_k}{\eta} - \frac{1}{L} (W-2) \sum w_{kk} u_k u^k \\ &\quad + \frac{1}{L} W v - \frac{2}{L} \sum w_{kk}^2 u_k u^k \\ &\leq \frac{1}{L} Q. \end{aligned}$$

Therefore we have

$$(3.39) \quad \sum v_{\alpha\alpha} \leq \frac{1}{L} Q.$$

It then follows from (3.38)

$$(3.40) \quad \begin{aligned} \mathfrak{R}_\varphi &\geq \frac{2L^2}{\varepsilon^2} \sum w_{ii} \frac{\nabla_i(g \circ T)}{g \circ T} u^i \\ &\quad - (n+6-W) \frac{L^2}{\varepsilon^2} \sum \frac{u^i \nabla_i(g \circ T)}{g \circ T} \\ &\quad - \sum \frac{\nabla_{\alpha\alpha}^2(g \circ T)}{g \circ T} - \frac{L}{\varepsilon^2} Q. \end{aligned}$$

Now we compute  $\nabla_\alpha(g \circ T)$  and  $\sum \nabla_{\alpha\alpha}^2(g \circ T)$ . By (2.16) we have

$$(3.41) \quad \nabla_\alpha(g \circ T) = g_k T_\alpha^k = g_\alpha w_{\alpha\alpha}.$$

By differentiating (2.16), we have

$$\begin{aligned} \sum \nabla_{\alpha\alpha}^2(g \circ T) &= \sum g_{kl} T_\alpha^k T_\alpha^l + \sum g_k \nabla_\alpha T_\alpha^k \\ &= \sum g_k A^{kl} w_{l\alpha;\alpha} + \sum g_{\alpha\alpha} w_{\alpha\alpha}^2 - \sum g_k A_{k\alpha;\alpha} w_{\alpha\alpha}. \end{aligned}$$

Recalling that  $A^{kl} = \delta_{kl}$  at  $x_0$ , we have

$$A^{kl} w_{l\alpha;\alpha} = w_{k\alpha;\alpha} = w_{\alpha\alpha;k} + A_{k\alpha;\alpha} - A_{\alpha\alpha;k}.$$

By (3.30),

$$\begin{aligned} \sum \nabla_{\alpha\alpha}^2(g \circ T) &= \sum g_k w_{\alpha\alpha;k} + \sum g_k (A_{k\alpha;\alpha} - A_{\alpha\alpha;k}) \\ &\quad + \sum g_{\alpha\alpha} w_{\alpha\alpha}^2 - \sum g_k A_{k\alpha;\alpha} w_{\alpha\alpha} \\ &= -W \sum g_k \frac{\eta_k}{\eta} + \sum g_{\alpha\alpha} w_{\alpha\alpha}^2 + \sum g_k (A_{k\alpha;\alpha} - A_{\alpha\alpha;k}) \\ &\quad + \sum g_k (A_{\alpha\alpha;k} - A_{k\alpha;\alpha}) w_{\alpha\alpha}. \end{aligned}$$

Using the second formula in Lemma 1, we get

$$(3.42) \quad \begin{aligned} \sum \nabla_{\alpha\alpha}^2(g \circ T) &= -W \sum g_k \frac{\eta_k}{\eta} + \sum g_{\alpha\alpha} w_{\alpha\alpha}^2 + \frac{L^2}{\varepsilon^2} (W-n) \sum w_{kk} u^k g_k \\ &\quad + \frac{L^2}{\varepsilon^2} (W - \sum w_{\alpha\alpha}^2) \sum u^k g_k. \end{aligned}$$

Inserting (3.41) and (3.42) into (3.40), we then obtain

$$(3.43) \quad \begin{aligned} (g \circ T) \mathfrak{R}_\varphi &\geq W \sum g_k \frac{\eta_k}{\eta} + \frac{2L^2}{\varepsilon^2} \sum w_{kk}^2 g_k u^k \\ &\quad - \frac{6L^2}{\varepsilon^2} \sum w_{kk} u^k g_k - \sum g_{\alpha\alpha} w_{\alpha\alpha}^2 \\ &\quad - \frac{L^2}{\varepsilon^2} (W - \sum w_{\alpha\alpha}^2) \sum u^k g_k - \frac{L}{\varepsilon^2} Q. \end{aligned}$$

By (3.5),

$$\begin{aligned}
 (3.44) \quad \sum g_k \frac{\eta_k}{\eta} &= L \sum (\sigma^{ij} + \frac{L^2}{\varepsilon^2} u^i u^j) g_i \frac{\eta_j}{\eta} \\
 &= \frac{L}{\eta} (\sum \sigma^{ij} g_i \eta_j + \frac{L^2}{\varepsilon^2} (\sum g_i u^i) (\sum \eta_j u^j)).
 \end{aligned}$$

We have

$$\sum \sigma^{ij} g_i \eta_j = \langle Dg, D\eta \rangle \leq C,$$

where  $D$  is the normal derivative in  $\mathbb{R}^n$  and  $\langle \cdot, \cdot \rangle$  denotes the standard Euclidean metric. Similarly,  $\sum \sigma^{ij} u_i g_j = \sum u^j g_j$ ,  $\sum \sigma^{ij} u_i \eta_j = \sum u^j \eta_j$  and  $\sum \sigma^{ij} g_i g_j$  are all bounded by a universal constant  $C$ . Hence from (3.44),

$$(3.45) \quad \sum g_k \frac{\eta_k}{\eta} \geq -\frac{L}{\varepsilon^2} \frac{C}{\eta}.$$

Employing (3.9) and (3.20),

$$(3.46) \quad (u^k)^2 = \frac{\varepsilon^2}{L^3} u_k u^k \leq \frac{\varepsilon^2}{L^3},$$

for any given  $k$ . Using (3.10) then (3.9), we have

$$\sum \sigma^{ij} g_i g_j = \frac{1}{L} (\sum g_i^2 - \frac{L^3}{\varepsilon^2} (\sum u^i g_i)^2).$$

It implies

$$\begin{aligned}
 (3.47) \quad \sum g_i^2 &= L \sum \sigma^{ij} g_i g_j + \frac{L^3}{\varepsilon^2} (\sum u^i g_i)^2 \\
 &\leq C \frac{L^3}{\varepsilon^2}.
 \end{aligned}$$

By (3.46) and (3.47) it follows that  $|g_k u^k| \leq C$ . Hence

$$(3.48) \quad \sum w_{kk}^2 g_k u^k \geq -CW^2$$

and

$$(3.49) \quad \sum w_{kk} u^k g_k \leq CW.$$

Moreover, in view of (3.8) and (3.9),

$$\sigma_{ii} = L + \frac{L^3}{\varepsilon^2} u_i u^i \leq C \frac{L}{\varepsilon^2},$$

for any given  $i$ . Consequently,

$$(3.50) \quad \sum g_{\alpha\alpha} w_{\alpha\alpha}^2 \leq |D^2 g| \sum \sigma_{\alpha\alpha} w_{\alpha\alpha}^2 \leq C \frac{L}{\varepsilon^2} W^2.$$

By virtue of (3.45), (3.48), (3.49) and (3.50), we obtain from (3.43) that

$$\begin{aligned}\mathfrak{R}_\varphi &\geq -\frac{L}{\varepsilon^2(g \circ T)}Q \\ &\geq -\frac{L}{\varepsilon^2}Q.\end{aligned}$$

This completes the proof.  $\square$

By Lemma 3 and (3.34), we get, at  $x_0$ ,

$$\begin{aligned}0 &\geq \frac{L}{\varepsilon^2}\mathcal{W}\sum w_{ii}^2 - C\frac{L}{\varepsilon^2}\left(1 + \frac{W}{\eta} + W^2 + \frac{W}{\eta}\mathcal{W}\right) \\ &\geq \frac{L}{\varepsilon^2}\mathcal{W}\frac{W^2}{n} - C\frac{L}{\varepsilon^2}\left(1 + \frac{W}{\eta} + W^2 + \frac{W}{\eta}\mathcal{W}\right).\end{aligned}$$

Multiplying  $n\eta^2L$  to both sides of the above inequality, we obtain

$$\begin{aligned}(3.51) \quad 0 &\geq \frac{L^2}{\varepsilon^2}\mathcal{W}(H^2 - CH) - C\frac{L^2}{\varepsilon^2}(1 + H^2) \\ &\geq C\frac{L^2}{\varepsilon^2}\mathcal{W}H^2 - C\frac{L^2}{\varepsilon^2}(1 + H^2).\end{aligned}$$

Note that by (3.11),

$$(3.52) \quad \mathcal{W} \geq \sum_{k \geq 2} \frac{1}{\lambda_k} \geq \left(\prod_{k \geq 2} \frac{1}{\lambda_k}\right)^{\frac{1}{n-1}} \geq C\lambda_1^{\frac{1}{n-1}} \geq C\left(\frac{W}{n}\right)^{\frac{1}{n-1}},$$

where  $C$  is independent of  $\varepsilon$ . Hence from (3.51) we get

$$(3.53) \quad 0 \geq \frac{L^2}{\varepsilon^2}H^{2+\frac{1}{n-1}} - C\frac{L^2}{\varepsilon^2}(1 + H^2).$$

Hence  $H \leq C$  at  $x_0$  and this completes the proof of Theorem 1.

#### 4. A COUNTEREXAMPLE TO THE LIPSCHITZ REGULARITY

In the last section we proved that the eigenvalues of  $DT_\varepsilon$  are uniformly bounded. In this section we give an example to show that the  $T_\varepsilon$  is not uniformly Lipschitz continuous for small  $\varepsilon > 0$ , i.e., the matrix  $DT_\varepsilon$  is not uniformly bounded, even though the densities  $f$  and  $g$  are smooth and positive, and the domain  $\Omega^*$  is c-convex with respect to  $\Omega$ .

Our counterexample will be obtained by finding a choice of  $f$  and  $g$  such that the monotonic optimal transport  $T_0$  between them is not Lipschitz continuous. Even if we said that the convergence  $T_\varepsilon \rightarrow T_0$  is not straightforward, we can prove that a uniform Lipschitz bound on  $T_\varepsilon$  would imply such a convergence, and hence the same bound on  $T_0$ . Hence, if  $T_0$  is not Lipschitz, then  $T_\varepsilon$  cannot be uniformly Lipschitz.

**Lemma 4.** *Suppose that the sequence of transports  $T_\varepsilon$  is uniformly Lipschitz. Then the whole family  $T_\varepsilon$  converges uniformly as  $\varepsilon \rightarrow 0$  to the unique monotonic optimal transport for the cost  $|x - y|$ , which will be Lipschitz with the same Lipschitz constant.*

*Proof.* By Ascoli-Arzelà's Theorem, the uniform Lipschitz bound implies the existence of a uniform limit up to subsequences. Obviously this limit map  $T$  will be optimal for the limit problem, i.e. the Monge problem for cost  $c(x, y) = |x - y|$  and will share the same Lipschitz constant as  $T_\varepsilon$ .

We only need to prove that  $T$  is monotonic along transport rays. Take  $L_\varepsilon(x) = \sqrt{\varepsilon^2 + |T_\varepsilon(x) - x|^2}$ : these maps are also uniformly Lipschitz and converge uniformly to  $L(x) = |T_\varepsilon(x) - x|$ . Let us denote by  $u_\varepsilon$  the potentials for the approximated problems and by  $u$  the potential for the limit problem. Due to the uniqueness of the Kantorovich potential  $u$ , since all the functions  $u_\varepsilon$  are 1-Lipschitz, we have  $u_\varepsilon \rightarrow u$  uniformly. Moreover,  $Du_\varepsilon \rightharpoonup Du$  and the convergence is actually strong (in  $L^2$ , for instance) if restricted to the set  $Tu = \{|Du| = 1\}$  (as a consequence of  $|Du_\varepsilon| \leq 1$ , which implies that we also have  $\int_{Tu} |Du_\varepsilon|^2 \rightarrow \int_{Tu} |Du|^2$ : this turns weak convergence into strong, and hence also implies pointwise, convergence).

The monotonicity of  $T$  is proven if one proves  $DL \cdot Du \leq 1$ , since the direction of the transport rays is that of  $-Du$ . This inequality is needed on the set of interior points of transport rays, which are exactly points where  $|Du| = 1$ . On these points we can use the weak convergence  $DL_\varepsilon \rightharpoonup DL$  (weakly-\* in  $L^\infty$ ) and the strong convergence  $Du_\varepsilon \rightarrow Du$ , which means that it is enough to get  $DL_\varepsilon \cdot Du_\varepsilon \leq 1$ , and then pass the inequality to the limit. This is the point where we use the uniform Lipschitz bound on  $T_\varepsilon$ : without such a bound we could not have the suitable weak convergence of  $DL_\varepsilon$ .

In order to estimate  $DL_\varepsilon$ , we use (2.8) and (2.4). We come back to the notation without the index  $\varepsilon$ , and write  $DL$ , thus getting

$$DL \cdot Du = -D_i u (T_j^i - \delta_{ij}) D_j u = L D_i u D_{ij}^2 u D_j u + |Du|^2 DL \cdot Du.$$

Then, we use (2.10) and (2.12) and the positivity of the matrix  $w_{ij}$ , to get

$$L D_i u D_{ij}^2 u D_j u \leq |Du|^2 (1 - |Du|^2).$$

This implies

$$(1 - |Du|^2) DL \cdot Du \leq |Du|^2 (1 - |Du|^2),$$

which provides  $DL \cdot Du \leq |Du|^2 \leq 1$  (notice that, for fixed  $\varepsilon > 0$ , the norm of the gradient  $|Du|$  is strictly less than 1, which allows to divide by  $1 - |Du|^2$ ).  $\square$

To construct the counterexample where  $T_0$  is not Lipschitz, our idea is as follows. Let

$$(4.1) \quad \ell_a = \{(x, y) \text{ in } \mathbb{R}^2 \mid y = \sqrt{a}(x + 2 + a), x \in [-2 - a, 1]\}$$

be a family of line segments  $\ell_a$ , where  $a \in [0, 1]$ . It is clear that the segments  $\ell_a$  do not intersect with each other and  $\cup_{a \in [0, 1]} \ell_a = \Delta_{ABC}$ , where  $\Delta_{ABC}$  denotes the triangle with vertices  $A = (-3, 0)$ ,  $B = (1, 4)$  and  $C = (1, 0)$ . Let

$$\begin{aligned} f &\equiv 1, \\ g &= 1 + \frac{1}{4}x + \eta(y) \end{aligned}$$

be two densities on  $\Delta_{ABC}$ . We first show that there exists a smooth positive function  $\eta$  such that  $f, g$  satisfy the mass balance condition

$$(4.2) \quad \int_{\Delta_{P_a C Q_a}} f = \int_{\Delta_{P_a C Q_a}} g, \text{ for all } a \in [0, 1].$$

Here  $P_a = (-2 - a, 0)$  and  $Q_a = (1, (3 + a)\sqrt{a})$  are the endpoints of  $\ell_a$ . We then prove that there is a Lipschitz function  $u$ , which is the potential function to Monge's problem in  $\Delta_{ABC}$ , with the densities  $f, g$  given above. By (4.2) we can construct a measure preserving mapping  $T_0$ , which pushes the density  $f$  to the density  $g$ , with  $\{\ell_a\}$  as its transfer rays. Using the potential  $u$  and the duality we show that  $T_0$  is the optimal mapping of Monge's problem. By reflection in the  $x$ -axis, we extend  $T_0$  to the triangle  $\Delta_{ABB'}$ , where  $B' = (1, -4)$  is the reflection of  $B$ . Then  $T_0$  is not Lipschitz at the interior point  $(-2, 0)$ .

**Lemma 5.** *There exists a smooth positive function  $\eta$ , such that (4.2) holds. This function satisfies  $\eta(y) = O(y^2)$  as  $y \rightarrow 0$ .*

*Proof.* By direct computation,

$$\begin{aligned} \int_{\Delta_{P_a C Q_a}} f &= \frac{1}{2}\sqrt{a}(3 + a)^2, \\ \int_{\Delta_{P_a C Q_a}} g &= \int_{-2-a}^1 \int_0^{\sqrt{a}(x+2+a)} \left(1 + \frac{1}{4}x + \eta(y)\right) dy dx \\ &= \frac{\sqrt{a}}{24}(3 + a)^2(12 - a) + \int_{-2-a}^1 \int_0^{\sqrt{a}(x+2+a)} \eta(y) dy dx. \end{aligned}$$

In order that (4.2) holds, we need

$$(4.3) \quad \frac{1}{24}a^{3/2}(3 + a)^2 = \int_{-2-a}^1 \int_0^{\sqrt{a}(x+2+a)} \eta(y) dy dx.$$

Differentiating (4.3) with respect to  $a$ , we have

$$\frac{a}{24} (9 + 7a) (3 + a) = \int_{-2-a}^1 (x + 2 + 3a) \eta(\sqrt{a}(x + 2 + a)) dx$$

which is equivalent to

$$(4.4) \quad \frac{a^2}{24} (9 + 7a) (3 + a) = \int_0^{(3+a)\sqrt{a}} (t + 2a\sqrt{a}) \eta(t) dt.$$

In order to find  $\eta$  satisfying (4.3) for all  $a \in [0, 1]$ , we only need to solve (4.4), since the equality in (4.3) is true for  $a = 0$ .

Let us introduce

$$(4.5) \quad y = (3 + a) \sqrt{a}.$$

It is clear that  $y$  is a strictly increasing function of  $a$ . Let  $a(y) = O(y^2)$  be the inverse function of (4.5). Differentiating (4.4) in  $y$  and using  $a_y = \frac{2\sqrt{a}}{3(a+1)}$ , we obtain

$$\frac{\sqrt{a}}{36} (27 + 45a + 14a^2) = \frac{3(1+a)^2}{2\sqrt{a}} \eta(y) + \int_0^y \eta(t) dt,$$

Taking derivative again, we obtain

$$(4.6) \quad \eta'(y) + \frac{q(a(y))}{y} \eta(y) = yp(a(y)),$$

where

$$\begin{aligned} q(a) &= \frac{(5a-1)(3+a)}{3(1+a)^2}, \\ p(a) &= \frac{27 + 135a + 70a^2}{162(1+a)^3(3+a)}. \end{aligned}$$

Solving (4.6), one finds an explicit formula for  $\eta$ :

$$(4.7) \quad \eta(y) = \int_0^y t p(a(t)) \exp\left(-\int_t^y \frac{q(a(\tau))}{\tau} d\tau\right) dt.$$

It is clear that

$$-1 \leq q(a(y)) \leq 0 \text{ if } |y| \ll 1.$$

Hence

$$\begin{aligned} 0 \leq \eta(y) &\leq C \int_0^y t \exp\left(\int_t^y \frac{1}{\tau} d\tau\right) dt \\ &\leq Cy^2. \end{aligned}$$

From (4.7) it follows that

$$\begin{aligned}
 (4.8) \quad \eta(y) &= \int_0^y t p(a(t)) \exp\left(-\frac{1}{2} \int_{a(t)}^{a(y)} \frac{5a-1}{a(1+a)} da\right) dt \\
 &= \frac{\sqrt{a}}{324(a+1)^3} \int_0^a \frac{3(s+1)(27+135s+70s^2)}{\sqrt{s}} ds \\
 &= \frac{a(10a^3+41a^2+54a+27)}{54(a+1)^3}.
 \end{aligned}$$

In the last two equalities,  $a$  is the function of  $y$  determined by (4.5). Therefore  $\eta$  is positive and smooth and satisfies the required conditions.  $\square$

**Remark 1.** From (4.5), we can explicitly write

$$a(y) = h(y) + \frac{1}{h(y)} - 2,$$

where

$$h(y) = \sqrt[3]{\sqrt{\frac{1}{4}y^4 + y^2} + \frac{1}{2}y^2 + 1}.$$

It is clear that  $a(y)$  is a smooth even function.

**Lemma 6.** There exists a function  $u : \Delta_{ABC} \rightarrow \mathbb{R}$  satisfying

$$|u(p) - u(q)| \leq |p - q|, \quad \forall p, q \in \Delta_{ABC},$$

and equality holds if and only if both  $p$  and  $q$  lie on a common segment  $\ell_a$ .

*Proof.* We will construct a function  $u : \Delta_{ABC} \rightarrow \mathbb{R}$ , which decreases linearly along all  $\ell_a$ .

For  $(x, y) \in \Delta_{ABC}$ , let  $a = a(x, y)$  be the solution of the equation

$$(4.9) \quad y = \sqrt{a}(a + 2 + x).$$

Hence  $(x, y) \in \ell_a$ . Differentiating (4.9) with respect to  $x$  and  $y$  respectively, we get

$$(4.10) \quad 0 = \frac{a_x}{2a}y + \sqrt{a}(a_x + 1)$$

and

$$(4.11) \quad 1 = \frac{a_y}{2a}y + \sqrt{a}a_y.$$

It follows by (4.10) and (4.11) that

$$(4.12) \quad a_y + \frac{a_x}{\sqrt{a}} = 0$$

provided  $a(x, y) \neq 0$ .



On the other hand, for  $(x, y) \in \Delta_{ABC}$ , the direction vector of the line segment  $\ell_a$  passing through  $(x, y)$  is given by

$$(4.13) \quad \nu(x, y) = (\nu_1, \nu_2) = -\frac{(1, \sqrt{a(x, y)})}{\sqrt{1 + a(x, y)}}.$$

Hence, by (4.12),

$$(4.14) \quad \partial_y \nu_1 - \partial_x \nu_2 = \frac{1}{2(1+a)^{3/2}} \left( a_y + \frac{a_x}{\sqrt{a}} \right) = 0,$$

provided  $a(x, y) \neq 0$ .

Fix a point  $P = (-2, 1)$ . Let

$$\begin{aligned} \gamma(t) &= \gamma_X(t) \\ &= (t(x+2) - 2, 1 - t(1-y)), \end{aligned}$$

$t \in [0, 1]$ . Then  $\gamma$  is the segment joining  $P$  and  $X = (x, y) \in \Delta_{ABC}$ . Set

$$u(x, y) = (x+2) \int_0^1 \nu_1(\gamma(t)) dt + (y-1) \int_0^1 \nu_2(\gamma(t)) dt.$$

We claim that  $u$  satisfies

$$(4.15) \quad Du(x, y) = \nu(x, y) \text{ on all segments } \ell_a.$$

Indeed, for any point  $X_0 = (x_0, y_0) \in \Delta_{ABC}$  with  $a(x_0, y_0) \neq 0$ , by (4.14) we have

$$\begin{aligned} (4.16) \quad u_x(x_0, y_0) &= \int_0^1 \nu_1(\gamma_0(t)) dt + (x_0+2) \int_0^1 t \partial_x \nu_1(\gamma_0(t)) dt \\ &\quad + (y_0-1) \int_0^1 t \partial_x \nu_2(\gamma_0(t)) dt \\ &= \int_0^1 \nu_1(\gamma_0(t)) dt + \int_0^1 t \frac{d}{dt} \nu_1(\gamma_0(t)) dt \\ &= \int_0^1 \frac{d}{dt} (t \nu_1(\gamma_0(t))) dt = \nu_1(x_0, y_0), \end{aligned}$$

where  $\gamma_0 = \gamma_{X_0}$  and we used  $\partial_x \nu_2 = \partial_y \nu_1$ . Similarly, we have

$$u_y(x_0, y_0) = \nu_2(x_0, y_0).$$

Taking limit, we see that (4.15) also holds on the segment  $\ell_a \mid_{a=0}$ .

As  $\nu$  is a unit vector, hence from (4.15) we have

$$|u(p) - u(q)| \leq |p - q|, \quad \forall p, q \in \Delta_{ABC},$$

and equality holds if and only if both  $p$  and  $q$  lie on a common segment  $\ell_a$ . This completes the proof.  $\square$

As in [8, 32], one can show by Lemma 5 that there is a unique measure preserving map  $T_0$  from  $(f, \Delta_{ABC})$  to  $(g, \Delta_{ABC})$  such that  $T_0(p)$  and  $p$  lie in a common  $\ell_a$  for all  $p \in \Delta_{ABC}$ , and satisfies the monotonicity condition

$$(T_0(p) - T_0(q)) \cdot (p - q) \geq 0 \quad \forall p, q \in \ell_a.$$

With the help of and Lemma 6, we prove that this  $T_0$  is indeed optimal. This fact is classical in optimal transport theory, but we show it in details for the sake of completeness.

**Lemma 7.**  *$T_0$  is an optimal mapping in the Monge mass transportation problem from  $(f, \Delta_{ABC})$  to  $(g, \Delta_{ABC})$ .*

*Proof.* Recall that the total cost functional is given by

$$\mathcal{C}(s) = \int_{\Delta_{ABC}} f(z) |z - s(z)| dz,$$

where  $s \in \mathcal{S}$ , the set of measure preserving maps from  $(f, \Delta_{ABC})$  to  $(g, \Delta_{ABC})$ ; and the Kantorovich functional is defined as

$$I(\psi, \varphi) = \int_{\Delta_{ABC}} f\psi + \int_{\Delta_{ABC}} g\varphi,$$

where  $(\psi, \varphi)$  are function pairs in the set

$$\mathcal{K} = \{\psi(x) + \varphi(y) \leq |x - y| \quad \forall x, y \in \Delta_{ABC}\}.$$

For all  $s \in \mathcal{S}$  and  $(\psi, \varphi) \in \mathcal{K}$ , we have

$$\begin{aligned} (4.17) \quad I(\psi, \varphi) &= \int_{\Delta_{ABC}} f(z) \psi(z) dz + \int_{\Delta_{ABC}} f(z) \varphi(s(z)) dz \\ &\leq \int_{\Delta_{ABC}} f(z) |z - s(z)| dz \\ &= \mathcal{C}(s). \end{aligned}$$

That is

$$\sup_{\mathcal{K}} I(\psi, \varphi) \leq \inf_{\mathcal{S}} \mathcal{C}(s).$$

Let  $u$  be the function constructed in the proof of Lemma 6, and let  $v = -u$ . Then we have  $(u, v) \in \mathcal{K}$ . As  $T_0(p)$  and  $p$  lie on the same line segment, Lemma 6 implies

$$u(z) - u(T_0(z)) = |z - T_0(z)|.$$

So the inequality in (4.17) becomes equality provided  $(\psi, \varphi) = (u, v)$  and  $s = T_0$ . Therefore

$$\mathcal{C}(T_0) = I(u, v) \leq \sup_{\mathcal{K}} I(\psi, \varphi) \leq \inf_{\mathcal{S}} \mathcal{C}(s).$$

Hence  $T_0$  is optimal and the segments  $\ell_a$  are transfer rays of Monge's problem.  $\square$

Let  $B' = (1, -4)$  be the reflection of the point  $B$  in the  $x$ -axis and let  $\Omega = \Omega^* = \Delta_{ABB'}$ . Extend the functions  $f, g$  to  $\Omega$  such that they are symmetric with respect to the  $x$ -axis. From the proof of Lemma 5, one sees that  $f, g$  are smooth and satisfy the mass balance condition (1.2). The fact that  $\eta$  is quadratic close to 0 shows that it can be reflected as a  $C^2$  function, and Remark 1 shows that it is indeed smooth. It is also known [26] that  $\Omega$  and  $\Omega^*$  are  $c$ -convex with respect to each other (for the cost function  $c_\varepsilon$ ,  $0 \leq \varepsilon \leq 1$ ).

Also extend  $T_0$  to  $\Omega$  so that it is symmetric with respect to the  $x$ -axis. By the uniqueness of monotone optimal mappings [17],  $T_0$  is an optimal mapping of Monge's problem from  $(f, \Omega)$  to  $(g, \Omega)$ .

We claim that  $T_0$  is not Lipschitz continuous at the point  $q_0 = (-2, 0)$ . Let  $D_{a,\delta}$  be the strip in  $\Delta_{ABC}$  between the segments  $\ell_a$  and  $\ell_{a+\delta}$ , and let  $q_\sigma = (-2, \sigma)$  be the intersection of  $\ell_a$  with the line  $\{x = -2\}$ , where  $\delta, \sigma > 0$  are constants. Let  $T_0(q_\sigma) = (x_\sigma, y_\sigma)$ . As  $T_0$  is measure preserving, we have (see the construction of the optimal mappings in [8, 32])

$$\lim_{\delta \rightarrow 0} \frac{1}{\delta} \int_{D_{a,\delta} \cap \{x < -2\}} f(x, y) dx dy = \lim_{\delta \rightarrow 0} \frac{1}{\delta} \int_{D_{a,\delta} \cap \{x < x_\sigma\}} g(x, y) dx dy.$$

That is

$$\int_{-2-a}^{-2} (x+2+3a) dx = \int_{-2-a}^{x_\sigma} (x+2+3a) \left(1 + \frac{1}{4}x + \eta(\sqrt{a}(x+2+a))\right) dx.$$

Making the change  $t = 2 + a + x$ , we obtain

$$\int_0^a (t+2a) dt = \int_0^{x_\sigma+2+a} (t+2a) \left(\frac{1}{2} + \frac{t-a}{4} + \eta(\sqrt{a}t)\right) dx$$

Since both  $(t-a)$  and  $\eta(\sqrt{a}t)$  tend to 0 when  $t, a \rightarrow 0$  (recall that  $\eta(t) = O(t^2)$ ), they are negligible in front of the constant  $\frac{1}{2}$ . This implies that, for small  $a$ , we should have

$$(4.18) \quad x_\sigma \geq -2 + (\sqrt{5} - 2)a.$$

Indeed, either  $x_\sigma + 2$  does not tend to 0, in which case (4.18) is satisfied, or it tends to 0, in which case we can write, for small  $a$ ,

$$\int_0^a (t+2a) dt \leq \int_0^{x_\sigma+2+a} \frac{3}{4}(t+2a) dx.$$

Computing these integrals explicitly we get exactly the inequality (4.18).

On the other hand, by (4.1), we have  $\sigma = a^{3/2}$ . Note that  $x(0) = -2$ . Hence

$$(4.19) \quad \lim_{\sigma \rightarrow 0+} \frac{x(\sigma) - x(0)}{\sigma} \geq \frac{1}{4} \lim_{\sigma \rightarrow 0+} a^{-1/2} = \infty.$$

Our claim follows.

As  $q_0 = (-2, 0)$  is an interior point of  $\Delta_{ABB'}$ , we have thus constructed positive, smooth densities  $f, g$ , and  $c$ -convex domains  $\Omega = \Omega^* = \Delta_{ABB'}$ , such that the associated optimal mapping  $T_0$  is not Lipschitz at interior points.

As the triangle  $\Delta_{ABB'}$  is  $c$ -convex with respect to each other, the optimal mapping  $T_\varepsilon$  is smooth [26]. By Lemma 4, one has  $T_0 = \lim_{\varepsilon \rightarrow 0} T_\varepsilon$ , and the above example shows that  $T_\varepsilon$  is not locally, uniformly Lipschitz continuous as  $\varepsilon \rightarrow 0$ .

## 5. APPLICATIONS AND PERSPECTIVES

The regularity problem for the Monge cost is very natural in transport theory and very difficult. For the moment, even the implication  $f, g \in C^\infty \Rightarrow T_0 \in C^0$  in a convex domain is completely open. The transport  $T_0$ , among all the optimal transports for the cost  $|x - y|$  (for which there is no uniqueness), is likely to be the most regular and the easiest to approximate.

The present paper presented a strategy inspired by the previous results introduced in [26] to get Lipschitz bounds, i.e.  $L^\infty$  bounds on the Jacobian. Yet, it only allows for some partial bounds, and the counter-example of Section 4 shows that a Lipschitz result is not possible. However, in the same counter-example, the monotonic transport  $T_0$  is a continuous map, and the point where a non-Lipschitz behavior is observed shows anyway the behavior of a  $C^{0, \frac{2}{3}}$  map. Thus, it is still possible to hope for continuous, or even Hölder, regularity results on  $T_0$ .

We stress that these results could also be applied to the regularity of the transport density. The transport density is a notion which is specifically associated to the transport problem for the Monge cost (see [17]): it is a measure  $\sigma$  which satisfies

$$(5.1) \quad \begin{cases} \operatorname{div} \cdot (\sigma Du) = f - g & \text{in } \Omega \\ |Du| \leq 1 & \text{in } \Omega, \\ |Du| = 1 & \text{a.e. on } \sigma > 0, \end{cases}$$

together with the Kantorovich potential  $u$ .

Several weak regularity results have been established, starting from the absolute continuity of  $\sigma$  if either  $f$  or  $g$  are absolutely continuous, till the  $L^p$  estimates  $f, g \in L^p \Rightarrow \sigma \in L^p$  (see [17, 1, 13, 14, 28]).

An explicit formula for  $\sigma$  in terms of optimal transport plans or maps is available (we will not develop it here, see [1]) and most possible strategies for the regularity of the transport density need some continuity of the corresponding optimal transport. Yet, one of the advantages of working with  $\sigma$  is that any optimal transport  $T$  produces the same density  $\sigma$ . This allows for choosing the most regular one, for instance  $T_0$ , but requires

anyway some regularity on it. Here is where our analysis comes into play (without, unfortunately, providing any exploitable result). But there are other features of the transport density that one could take advantage of: from the fact that it only depends on the difference  $f - g$ , one can decide to add any common density to both measures. For instance, if  $f$  and  $g$  are smooth densities with compact support on  $\mathbb{R}^n$ , it is always possible to add common background measure on a same convex domain  $\Omega$  including both the supports.  $\Omega$  can be chosen as smooth as we want, and we can for instance take  $\Omega$  to be a ball. Also, one can add another common density to  $f$  and  $g$  so as to get  $g = 1$ . This last trick allows to avoid some of the tedious computations of Section 3, since in this case  $g(T)$  has not to be differentiated.

In any case, even with these simplifications, the continuity result is not available for the moments. Possible perspectives of the current research involve the use of these partial estimates to prove continuity.

Among the possible strategies

- Use the bounds on  $DT_\varepsilon$  to get estimates on the directions of the transport rays for the limit problem, and use them to estimate how much the disintegrations of  $f$  and  $g$  vary according to the rays. Using the fact that the monotonic optimal transport (in one dimension) continuously depends on the measures, one can hope for the continuity of  $T_0$ .
- Use the fact that the bound on  $W$  gives an  $L^\infty$  bound on  $\operatorname{div}(LDu)$  and, since  $L$  depends on  $|Du|$ , one faces a highly non-linear and highly degenerate elliptic PDE where the goal would be to get uniform continuity results on  $LDu$ . This recalls what has been recently done in very degenerate elliptic PDEs for traffic applications (see [29, 11]), but seems (much) harder because  $LDu$  is not a uniformly continuous function of  $Du$ .
- Write down some elliptic PDEs solved by some scalar quantities associated to  $T_\varepsilon$ , for instance by  $L$ , and use the bounds on the matrices  $A$  and  $w$  that have been proven here in order to apply De Giorgi-Moser arguments (or their wider generalizations, see [15] for a complete framework). Should it work, this would give Hölder continuity. Unfortunately, our attempts have not given any useful PDE so far.

All in all, up to the two-dimensional result of [19] (which requires disjoint and convex supports), the search for continuous optimal transports for the original cost of Monge is still widely open.

## REFERENCES

- [1] Ambrosio, L.: Lecture notes on optimal transport problems, *Mathematical Aspects of Evolving Interfaces*, pp.1-52, *Lecture Notes in Math.*, vol.1812, (2003).
- [2] Ambrosio, L., Rigot, S.: Optimal mass transportation in the Heisenberg group, *J. Funct. Anal.*, 208, 261-301 (2004).
- [3] Ambrosio, L., Kirchheim, B., Pratelli, A.: Existence of optimal transport maps for crystalline norms, *Duke Math. J.*, 125, 207–241 (2004).
- [4] Brenier, Y., Polar factorization and monotone rearrangement of vector-valued functions, *Comm. Pure Appl. Math.* 44, 375–417 (1991).
- [5] Caffarelli, L.: The regularity of mappings with a convex potential, *J. Amer. Math. Soc.*, **5**, 99-104 (1992).
- [6] Caffarelli, L.: Boundary regularity of maps with convex potentials II, *Ann. of Math.* **144**, 453–496 (1996).
- [7] Caffarelli, L.: Allocation maps with general cost functions, in *Partial Differential Equations and Applications* (P. Marcellini, G. Talenti, and E. Vesintini eds), *Lecture Notes in Pure and Appl. Math.*, 177, 29-35 (1996).
- [8] Caffarelli, L., Feldman, M., McCann, R.J.: Constructing optimal maps for Monge’s transport problem as a limit of strictly convex costs, *J. Amer. Math. Soc.* 15, 1-26 (2002).
- [9] Champion, T., De Pascale, L.: The Monge problem for strictly convex norms in  $\mathbf{R}^d$ , *J. Eur. Math. Soc.* 12 (2010), 1355–1369.
- [10] Champion, T., De Pascale, L.: The Monge problem in  $\mathbf{R}^d$ , *Duke Math. J.*, 157, 551-572 (2011).
- [11] M. Colombo, A. Figalli Regularity results for very degenerate elliptic equations, preprint, <http://cvgmt.sns.it/paper/1996/>, 2012.
- [12] Delanoë, P., Ge, Y.: Regularity of optimal transport on compact, locally nearly spherical, manifolds, *J. Reine Angew. Math.*, 646, 65-115 (2010).
- [13] L. De Pascale, L. C. Evans and A. Pratelli, Integral estimates for transport densities, *Bull. of the London Math. Soc.* 36, n. 3, pp. 383–385, 2004.
- [14] L. De Pascale and A. Pratelli, Sharp summability for Monge Transport density via Interpolation, *ESAIM Control Optim. Calc. Var.* 10, n. 4, pp. 549–552, 2004.
- [15] E. Di Benedetto, *Degenerate Parabolic Equations*, Universitext, Springer-Verlag, 1993.
- [16] Evans, L.C., Gangbo, W.: Differential equations methods for the Monge-Kantorovich mass transfer problem, *Mem. Amer. Math. Soc.*, 653, 1999.
- [17] Feldman, M., McCann, R.J.: Uniqueness and transport density in Monge’s mass transportation problem, *Cal. Var. PDE*, 15, 81-113 (2002).
- [18] Figalli, A., Rifford, L., Villani, C.: Nearly round spheres look convex, *Amer. J. Math.*, 134, 109-139 (2012).
- [19] Fragalà, I. Gelli, M.S., Pratelli, A.: Continuity of an optimal transport in Monge problem, *J. Math. Pures Appl.*, 84, 1261-1294 (2005).
- [20] Gangbo, W., McCann, R.J.: Optimal maps in Monge’s mass transport problem, *C.R. Acad. Sci. Paris, Series I, Math.*, 321, 1653-1658 (1995).
- [21] Kantorovich, L., On the transfer of masses. *Dokl. Acad. Nauk. USSR*, (37), 7–8, 1942.
- [22] Kim, Y.-H., McCann, R. J.: Continuity, curvature, and the general covariance of optimal transportation, *J. Eur. Math. Soc. (JEMS)* 12, 1009–1040 (2010).
- [23] Liu, J.K., Trudinger, N.S.; Wang, X.-J., Interior  $C^{2,\alpha}$  regularity for potential functions in optimal transportation, *Comm. Partial Differential Equations* 35 165–184 (2010).
- [24] Loeper, G.: On the regularity of solutions of optimal transportation problems, *Acta Math.*, 202, 241-283 (2009).
- [25] Loeper, G., Villani, C.: Regularity of optimal transport in curved geometry: the nonfocal case, *Duke Math. J.*, 151, 431-485 (2010).

- [26] Ma, X.N., Trudinger, N.S., Wang, X.J.: Regularity of potential functions of the optimal transportation problem, Arch. Rational Mech. Anal., 177, 151-183 (2005).
- [27] Monge, G.: Memoire sur la Theorie des Déblais et des Remblais, In: Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la Même année, 1781, 666-740.
- [28] F. Santambrogio, Absolute continuity and summability of transport densities: simpler proofs and new estimates, Calc. Var. Par. Diff. Eq. 36, no. 3, 2009, 343-354.
- [29] F.Santambrogio, V.Vespri: Continuity in two dimensions for a very degenerate elliptic equation. Nonlinear Anal. 73 (2010), 3832-3841.
- [30] Sudakov, V. N.: Geometric problems in the theory of infinite dimensional probability distributions, Proc. Steklov Inst., 141, 1-178 (1979).
- [31] Villani, C.: Optimal transport, old and new, Springer-Verlag, Berlin, 2009.
- [32] Trudinger, N.S., Wang, X.J.: On the Monge mass transfer problem, Calc. Var. PDE, 13, 19-31 (2001).
- [33] Trudinger, N.S., Wang, X.J.: On the second boundary value problem for Monge-Ampère type equations and optimal transportation, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 8 143-174 (2009).

CENTRE FOR MATHEMATICS AND ITS APPLICATIONS, AUSTRALIAN NATIONAL UNIVERSITY, CANBERRA, ACT 0200, AUSTRALIA.

*E-mail address:* qi-rui.li@anu.edu.au

LABORATOIRE DE MATHÉMATIQUES D'ORSAY, UNIVERSITÉ PARIS-SUD, 91405 ORSAY CEDEX, FRANCE

*E-mail address:* filippo.santambrogio@math.u-psud.fr

CENTRE FOR MATHEMATICS AND ITS APPLICATIONS, AUSTRALIAN NATIONAL UNIVERSITY, CANBERRA, ACT 0200, AUSTRALIA.

*E-mail address:* xu-jia.wang@anu.edu.au